MATH 350 Linear Algebra
Final Exam Review

Instructor: Will Chen

December 11, 2022

Here is a **non-comprehensive** overview of what we have covered. You should make sure you are familiar with these topics. Unless otherwise stated, you should be able to recite all definitions. You should be able to prove every theorem we've covered. As an exercise, when reading each theorem below, try to remember how it was proved. You should be able to answer all of the questions I pose below, and give justifications/proofs when relevant.

Caution: There are probably typos. If you think you've found one, please lmk ASAP! If you find a typo/mistake that could potentially cause mathematical confusion, you will receive **a bonus point on the exam.**

1 Exam 1 Review

1.1 Vector spaces, subspaces, and quotient spaces

Definition 1.1.1 (Vector space, abridged)**.** A vector space over a field F is a set V together with two operations $+ : V \times V \to V$ and $\cdot : F \times V \to V$ which satisfies the properties VS1-VS8 (see §1.2 in the book).

An element of a vector space is called a vector. We sometimes leave out the field F, and simply say "let V be a vector space". In this case it should be understood that V is a vector space over some field F. If we write "Let V, W be vector spaces", then it should be understood that they are vector spaces over the same field. You should know the definition of a field, but you do not need to memorize it for this exam.

Definition 1.1.2 (Fields)**.** A field F is a set with two operations $+ : F \times F \to F$ and $\cdot : F \times F \to F$, such that the following hold for all $a, b, c \in F$:

(F1) $a + b = b + a$ and $a \cdot b = b \cdot a$

(F2) $(a + b) + c = a + (b + c)$ and $(a \cdot b) \cdot c = a \cdot (b \cdot c)$

(F3) There exist distinct elements "0" and "1" in F such that

$$0 + a = a \quad \text{and} \quad 1 \cdot a = a$$

(F4) For each $a \in F$ and *nonzero* $b \in F$, there exist elements $c, d \in F$ such that

$$a + c = 0 \quad \text{and} \quad b \cdot d = 1$$

Here, c is called the negative of a, denoted "$-a$", and d is called the multiplicative inverse of b, denoted "b^{-1}" or "$1/b$".

(F5) $a \cdot (b + c) = a \cdot b + a \cdot c$

Definition 1.1.3 (Subspace)**.** A subspace of a vector space V is a subset $W \subset V$ satisfying:

(a) $0 \in W$

(b) For any $x, y \in W$, $x + y \in W$.

(c) For any $x \in W, a \in F$, $ax \in W$.

Recall that $A \subset B$ means A is a subset of B. This does not rule out the case $A = B$. If we want to say that A is a proper subset of B, we will write $A \subsetneq B$.

In class I mistakenly noted that the last two conditions in the definition imply the first. They do not! For example, the empty subset $\{\} \subset V$ satisfies the last two conditions, but does not satisfy the first. All three conditions are required!

As the course progresses you should be collecting a list of examples to refer to. For example, we have discussed F^n (where F is a field), spaces of functions (all functions, linear functions, continuous functions...etc), spaces of polynomials "$P_n(F)$" and "$P(F)$", spaces of $m \times n$ matrices...

Have an understanding of basic properties of subspaces of a vector space V.

- How many 0-dimensional subspaces are there?
- If $\dim V = n$, how many n-dimensional subspaces are there?
- What are the 1-dimensional subspaces in F^n?
- Is the intersection of 2 subspaces a subspace?
- Is the intersection of arbitrarily many subspaces a subspace?
- If $W_1, W_2 \subset V$ are subspaces, what is $W_1 + W_2$? What does $W_1 \oplus W_2$ mean? Are they subspaces?
- What are other ways of constructing subspaces?

You should be able to answer the questions above and give justifications when relevant.

Recall the definition of quotient spaces (also called coset spaces)

Definition 1.1.4. Let V be a vector space and $W \subset V$ a subspace. If $v \in V$, then

$$v + W := \{v + w \mid w \in W\}$$

is called the *coset of* W *containing* v. The *quotient space* V/W is defined to be the set

$$V/W := \{v + W \mid v \in V\}$$

equipped with the operations of addition and scalar multiplication:

$$(v + W) + (v' + W) = (v + v') + W \qquad a(v + W) = (av) + W \qquad \text{for all } v, v' \in V, a \in F$$

We showed in the homeworks that these operations are well defined and make V/W into a vector space.

- If $v, v' \in V$, when is $v + W = v' + W$? If $v \neq v'$, must $v + W \neq v' + W$?
- If $v + W \neq v' + W$, what is $(v + W) \cap (v' + W)$?
- If $(v + W) \cap (v' + W)$ is nonempty, must $v + W = v' + W$?

1.2 Linear combination, span and linear independence

Definition 1.2.1 (Linear combination)**.** If V is a vector space and $S \subset V$ a subset, then a linear combination of vectors in S is a vector of the form

$$a_1 s_1 + a_2 s_2 + \cdots + a_n s_n$$

for some collection of scalars $a_i \in F$ and $s_i \in S$.

Definition 1.2.2 (Span)**.** If V is a vector space and $S \subset V$ a subset, then $\operatorname{Span}(S)$ is the set of all linear combinations of vectors in S. We say that S spans V (or S generates V, or S is a spanning set for V) if $\operatorname{Span}(S) = V$.

Note that S doesn't have to be finite!

- What is the relationship between $S, \operatorname{Span}(S), V$?
- What is $\operatorname{Span}(V)$?
- If S, T are two subsets of V, how does $\operatorname{Span}(S \cup T)$ compare to $\operatorname{Span} S + \operatorname{Span} T$?

The following alternate characterization of span is often useful:

Theorem 1.2.3. *If V is a vector space and $S \subset V$ a subset, then* $\operatorname{Span}(S)$ *is the intersection of all subspaces containing S.*

The slogan here is that "$\operatorname{Span}(S)$ is the smallest subspace containing S". Here, by "slogan" I mean a phrase which is not precise, but is easy to remember and conveys the right intuition.

(a) What is not precise about the statement "$\operatorname{Span}(S)$ is the smallest subspace containing S"?

The concepts of span and linear independence are "dual" to each other. This is a philosophy to keep in mind. Whenever you see a theorem about one, think about what the corresponding statement would be for the other. Here are some examples:

Theorem 1.2.4 (Existence vs uniqueness)**.** *Let S be a subset of a vector space V. Then*

(a) S spans V if and only if every $v \in V$ is a linear combination of vectors in S.

(b) S is linearly independent if and only if any $v \in V$ which is a linear combination of vectors in S is a linear combination of vectors in S in a unique way*.*

The moral here is that span is essentially about "existence", whereas linear independence is about "uniqueness".

Theorem 1.2.5 (Permanence under subset vs superset)**.** *Let S be a subset of a vector space V. Then*

(a) If S spans V, and $S' \supset S$, then S' also spans V..

(b) If S is linearly independent and $S' \subset S$, then S' is also linearly independent.

The slogan is: "supersets of spanning sets are also spanning", and "subsets of linearly independent sets are also linearly independent".

Theorem 1.2.6 (Linear independence in terms of span)**.** *Let S be a subset of a vector space V, then S is linearly independent if and only if no element $s \in S$ is in the span of $S - \{s\}$.*

Here, the "dual" statement to the above can be considered to be itself.

1.3 Bases and dimension

The notions of basis and dimension go hand in hand.

Definition 1.3.1 (Basis)**.** A basis of a vector space V is a subset $S \subset V$ that is linearly independent and spans V.

The plural of basis is "bases". Know the replacement theorem:

Theorem 1.3.2 (Replacement theorem, Theorem 1.10 in the book)**.** *Let V be a vector space that is generated by a set G of size n. Let L be a linearly independent subset of V of size m. Then $m \leq n$, and there exists a subset H of G containing exactly $n - m$ vectors such that $L \cup H$ generates V.*

It follows from the replacement theorem that

Theorem 1.3.3 (All bases have the same size)**.** *Let V be a vector space with a finite spanning set. Then V has a finite basis, and any two bases for V have the same size.*

You should know how to prove this using the replacement theorem. Because of the theorem, the following definition makes sense:

Definition 1.3.4 (Dimension)**.** Let V be a vector space. If V has a finite spanning set, then the dimension of V is defined to be the size of any basis. If V doesn't have a finite spanning set, then its dimension is infinite.

You should know the two alternate characterizations of a basis:

Theorem 1.3.5. *Let V be a vector space.*

(a) A subset $S \subset V$ is a basis if and only if it is a minimal spanning set. In other words, S is a basis if and only if S spans V and any strict subset $S' \subsetneq S$ does not span V.

(b) A subset $S \subset V$ is a basis if and only if it is a maximal linearly independent set. In other words, S is linearly independent, and any strict superset $S' \supsetneq S$ is not linearly independent.

Slogan: A basis is a minimal spanning set. It is also a maximal linearly independent set.

If V is finite dimensional, then we have the following numerical criterion for a set to be a basis:

Theorem 1.3.6. *Let V be a vector space of dimension n. Then*

(a) A subset $S \subset V$ is a basis if and only if S is linearly independent and $|S| = n$.

(b) A subset $S \subset V$ is a basis if and only if S spans V and $|S| = n$.

Some questions:

- If V is an n-dimensional vector space. What are the possible sizes of linearly independent sets?
- If V is an n-dimensional vector space. What are the possible sizes of spanning sets?
- If $W \subset V$ is a subspace, what is the relation between $\dim W$ and $\dim V$?
- If $S \subset V$ is linearly independent, does there exist a basis for V that contains S? (I.e., can every linearly independent set be extended to a basis?)
- If $S \subset V$ is a basis, and $W \subset V$ a subset, must there exist a subset $S' \subset S$ such that S' is a basis for W?
- If $S \subset V$ spans V, must there exist $S' \subset S$ such that S' is a basis for V?
- If $S \subset V$ is linearly independent, is it a basis for a subspace of V?
- If $S \subset V$ spans V, is it a basis for a subspace of V?
- If $\mathbb{F}_p$ denotes a finite field with p elements and V is an n-dimensional vector space over $\mathbb{F}_p$, how many vectors are in V?

You should know the following theorem (i.e., know that it is true), but you do not need to know the proof (the proof basically involves the axiom of choice).

Theorem 1.3.7. *Every vector space has a maximal linearly independent set.*

- Does every vector space have a basis?

1.4 Linear transformations

Definition 1.4.1. A linear transformation is a function between vector spaces $f : V \to W$ such that

(a) For any $v, v' \in V$, $f(v + v') = f(v) + f(v')$, and

(b) For any $v \in V$, $a \in F$, $f(av) = af(v)$.

Sometimes for short we will simply say that $f : V \to W$ is *linear*. You should know that these properties imply that

$$f(a_1v_1 + a_2v_2 + \cdots + a_nv_n) = a_1f(v_1) + a_2f(v_2) + \cdots + a_nf(v_n)$$

for any $a_1, \dots, a_n \in F$ and $v_1, \dots, v_n \in F$. You should know how to prove this from the definition using induction.

As the course progresses, you should keep in mind a collection of examples of linear transformations. For example, differentiation, integration, rotation in $\mathbb{R}^2$, projection, reflection... There is also the zero linear transformation, and the identity linear transformation. Moreover, you can build linear transformations by adding linear transformations or multiplying by scalars. In other words,

Theorem 1.4.2. *Let V, W be vector spaces. The set of all linear transformations $V \to W$ is denoted $\mathcal{L}(V, W)$. Then $\mathcal{L}(V, W)$ is a vector space with the operations of addition and scalar multiplication.*

Definition 1.4.3. Let $f : V \to W$ be a linear transformation. The image of f, denoted $\operatorname{im}(f)$ and the kernel of f, denoted $\ker(f)$ are:

$$\begin{aligned} \operatorname{im}(f) &:= \{f(v) \mid v \in V\} = \{w \in W \mid w = f(v) \text{ for some } v \in V\} \\ \ker(f) &:= \{v \in V \mid f(v) = 0\} \end{aligned}$$

The image is also sometimes called the *range*. The kernel is sometimes also called the *nullspace*.

Theorem 1.4.4. *Let $f : V \to W$ be a linear transformation. Then $\operatorname{im}(f)$ is a subspace of W, and $\ker(f)$ is a subspace of V.*

Theorem 1.4.5 (Spans and linear transformations)**.** *Let $f : V \to W$ be a linear transformation. If $S \subset V$ spans V, then $f(S) := \{f(s) \mid s \in S\}$ spans $\operatorname{im}(f)$.*

Definition 1.4.6. The rank and nullity of f are defined as:

$$\begin{aligned} \operatorname{rank}(f) &:= \dim \operatorname{im}(f) \\ \operatorname{nullity}(f) &:= \dim \ker(f) \end{aligned}$$

Sometimes this is infinite.

When the domain V is finite dimensional, rank and nullity are finite and satisfy the

Theorem 1.4.7 (Dimension theorem, Theorem 2.3 in the book)**.** *Let $f : V \to W$ be linear. If V is finite dimensional, then*

$$\operatorname{rank}(f) + \operatorname{nullity}(f) = \dim(V)$$

- What happens if V is infinite dimensional? If V is infinite dimensional, could $\operatorname{rank}(f)$ and $\operatorname{nullity}(f)$ both be finite?

For linear transformations, we have the following useful characterizations of 1-1 and onto:

Theorem 1.4.8. *Let $f : V \to W$ be linear. Then*

(a) f is 1-1 if and only if $\ker(f) = 0$

(b) f is onto if and only if $\operatorname{rank}(f) = \dim W$.

A synonym for 1-1 is "injective". A synonym for onto is "surjective".

Theorem 1.4.9. *Let $f : V \to W$ be linear.*

(a) If f is injective, then for any linearly independent set $S \subset V$, $f(S)$ is also linearly independent. (Slogan: "f injective implies that f preserves linear independence")

(b) If f is surjective, then for any spanning set $S \subset V$, $f(S)$ also spans W. (Slogan: "f surjective implies that f preserves spanning sets")

More food for thought:

- Let $W \subset V$ be a subspace. Consider the map $p : V \to V/W$ defined by sending $v \mapsto v + W$. Is p linear? What is $\ker(p)$? What is $\mathrm{rank}(p)$? Is p surjective? injective?

The following theorem is crucially important for understanding the relation between matrices and linear transformations:

Theorem 1.4.10 (Theorem 2.6 in the book)**.** *Let V, W be vector spaces. Let $v_1, \ldots, v_n$ be a basis for V. Let $w_1, \ldots, w_n$ be arbitrary elements of W. Then there exists exactly one linear transformation $f : V \to W$ such that $T(v_i) = w_i$ for each i.*

In the theorem, the fact that there is at least one such f is due to $\{v_1, \ldots, v_n\}$ being linearly independent. The fact that there is at most one such f is due to $\{v_1, \ldots, v_n\}$ being a spanning set for V. Since $\{v_1, \ldots, v_n\}$ is a basis, it is both linearly independent and spanning, so there is a unique such f.

1.5 Matrices and linear transformations

Let V, W be vector spaces, and $\beta = (v_1, \ldots, v_n)$ an ordered basis for V. Recall that if A is a set and $n \geq 1$ an integer, then A^n denotes the n-fold cartesian product, whose elements are n-tuples $(a_1, \ldots, a_n)$ with each $a_i \in A$. If $|A| = m$, then $|A^n| = mn$.

Consider the map

$$\begin{array}{rcl} \Phi_\beta : \mathcal{L}(V, W) & \longrightarrow & W^n \\ f & \mapsto & (f(v_1), f(v_2), \ldots, f(v_n)) \end{array}$$

Theorem 1.4.10 implies:

Theorem 1.5.1. *The map Φ_β is a bijection (i.e., it is both 1-1 and onto).*

If $\gamma = (w_1, \ldots, w_m)$ is an ordered basis for W, then any $w \in W$ can be written as a linear combination of the w_i's *in a unique way.* I.e, for any $w \in W$, there exist uniquely determined coefficients $a_1, \ldots, a_m$ such that $w = a_1 w_1 + a_2 w_2 + \cdots + a_m w_m$. This tuple of coefficients is often viewed as an $m \times 1$ matrix, called the *coordinate vector of w relative to γ*, denoted

$$[w]_\gamma := \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix}$$

Let us write $([w]_\gamma)_i$ for the ith entry. In our notation above, we have $([w]_\gamma)_i = a_i$. Consider the map

$$\begin{array}{rcl} \Psi_\gamma : W^n & \longrightarrow & M_{m\times n}(F) \\ (x_1, \ldots, x_n) & \mapsto & A_{ij} \end{array}$$

where $A_{ij} := ([x_j]_\gamma)_i$. Then

Theorem 1.5.2. *The map Ψ_γ is a bijection.*

Composing Φ_β and Ψ_γ, we have

Theorem 1.5.3 (The matrix associated to a linear transformation)**.** *Let V, W be vector spaces. Let $\beta = (v_1, \ldots, v_n)$ be an ordered basis for V, $\gamma = (w_1, \ldots, w_m)$ an ordered basis for W. Then the map*

$$\Psi_\gamma \circ \Phi_\beta : \mathcal{L}(V, W) \longrightarrow M_{m\times n}(F)$$

is a bijective linear transformation. If $f \in \mathcal{L}(V, W)$, then the matrix of f relative to the bases β, γ is denoted

$$[f]_\beta^\gamma := \Psi_\gamma(\Phi_\beta(f))$$

You should understand this map. Note that $\Psi_\gamma \circ \Phi_\beta$ is also *linear*. You should be able to compute the matrix of a linear transformation (relative to bases of the domain and codomain). Given a matrix A, you should also be able to find a linear transformation whose matrix is A (relative to suitable bases).

Caution: Given a linear transformation $f : V \to W$, it does not make sense to speak of the matrix of f. To talk about the matrix of f, you need to choose a basis for V and a basis for W. Any two bases will do, and relative to any two choses bases, you can compute a matrix for f. The matrix will usually depend on the chosen bases. Similarly, given a matrix A, to define a linear transformation whose matrix is A, one needs to choose vector spaces V, W and appropriate bases for them. If A is $m \times n$, then a convenient choice is to choose $V = F^n, W = F^m$, and to choose the bases to be the standard bases.

- Can you give an example that shows that the matrix of a linear transformation depends on the choice of bases β, γ?

If $f : V \to W$ and $g : W \to Z$ are linear maps between vector spaces, then the composition $g \circ f$ is the map

$$\begin{array}{rcl} g \circ f : V & \longrightarrow & Z \\ v & \mapsto & g(f(v)) \end{array}$$

Note that you can only compose if the codomain of the first map matches the domain of the second.

Theorem 1.5.4. *With notation as above, $g \circ f$ is linear.*

Some basic properties of composition are as follows. All are easily proven from the definitions.

Theorem 1.5.5 (See Theorem 2.10 in the book for a more precise, but also more restrictive statement)**.** *Let f, g, h be linear transformations (between possibly lots of different vector spaces). Let* id *denote the identity linear transformation (of some vector space). Whenever composition makes sense, we have*

(a) $f \circ (g + h) = f \circ g + f \circ h$ *and* $(g + h) \circ f = g \circ f + h \circ f$

(b) $f \circ (g \circ h) = (f \circ g) \circ h$

(c) $f \circ \mathrm{id} = \mathrm{id} \circ f = f$

(d) $a(f \circ g) = (af) \circ g = f \circ (ag)$ *for any scalar $a \in F$.*

The fact that $\Psi_\gamma \circ \Phi_\beta$ is linear tells us that addition and scalar multiplication of linear transformations corresponds to addition and scalar multiplication of matrices. The next natural question is: what does composition of linear transformations correspond to in terms of matrices? In other words, if α, β, γ are bases for V, W, Z respectively, and $f : V \to W$, $g : W \to Z$ are linear, then what is the matrix $[g \circ f]_\alpha^\gamma$ in terms of f, g? The answer is:

Theorem 1.5.6 (Composition corresponds to matrix multiplication, Theorem 2.11 in the book)**.** *Let V, W, Z be finite dimensional vector spaces and let $f : V \to W$ and $g : W \to Z$ be linear transformations. Let α, β, γ be bases of V, W, Z respectively. Then*

$$[g \circ f]_\alpha^\gamma = [g]_\beta^\gamma [f]_\alpha^\beta$$

where the right hand side is the product of matrices.

Note that $[g]_\beta^\gamma, [f]_\alpha^\beta$ will generally have different sizes. Indeed, if $\dim V = n, \dim W = m, \dim Z = r$, then $[g]_\beta^\gamma$ will be $r \times m$, $[f]_\alpha^\beta$ will be $m \times n$, and $[g \circ f]_\beta^\gamma$ will be $r \times n$. You should know how to multiply matrices.

We did not discuss this final theorem in class, but the proof is quite short. It is a good exercise to try to prove it.

Theorem 1.5.7 (Theorem 2.14 in the book)**.** *Let V, W be finite dimensional vector spaces having ordered bases β and γ respectively. Let $f : V \to W$ be linear. Then for each $v \in V$, we have*

$$[f(v)]_\gamma = [f]_\beta^\gamma [v]_\beta.$$

Here, recall that $[f(v)]_\gamma$, $[v]_\beta$ are the coordinate vectors of $f(v)$ and v with respect to γ and β respectively.

Food for thought:

- Let A be an $m \times n$ matrix and B an $n \times p$ matrix. Let $(AB)_{*j}$ denote the jth column of AB. Does $(AB)_{*j}$ depend on every entry of the matrix B? Which entries does it depend on?
- Let A be an $m \times n$ matrix and B an $n \times p$ matrix. Let B_{*j} denote the jth column of B. Can you write B_{*j} as BC for some matrix C? (Hint: By considering sizes of the matrices, C would have to be a $p \times 1$ matrix)

2 Exam 2 Review

2.1 Invertibility, isomorphisms, and change of coordinates

Definition 2.1.1. A function $f : X \to Y$ is 1-1 if whenever $f(x) = f(x')$, then $x = x'$. The function f is onto if for every element $y \in Y$, there is an $x \in X$ such that $f(x) = y$. The function f is invertible if there exists a function $g : Y \to X$ such that $f \circ g = \mathrm{id}_Y$ and $g \circ f = \mathrm{id}_X$. In this case we will sometimes write g as f^{-1}.

The most basic result on invertibility is:

Theorem 2.1.2. *A function $f : X \to Y$ is invertible if and only if it is both 1-1 and onto.*

Since linear transformations are functions, this also applies to linear transformations. In this case, there is much more structure:

Theorem 2.1.3. *Let $f : V \to W$ be linear. Then f is 1-1 if and only if $N(f) = 0$, and f is onto if and only if* $\mathrm{rank}(f) = \dim W$.

Here, $N(f)$ is the null space of f (or equivalently the kernel of f.

Combining this result with the rank-nullity theorem gives even more information. For example, it tells us that

$$\ker(f) = 0 \iff \mathrm{rank}(f) = \dim V$$

If V, W are finite dimensional and β, γ are bases of V, W respectively, then rank, and hence nullity can be expressed in terms of the rank of $[f]_\beta^\gamma$. Namely,

Theorem 2.1.4. *Let $f : V \to W$ be linear, V, W finite dimensional, and β, γ be bases of V, W respectively. Then*

(a) $\mathrm{rank}(f) = $ *row-rank*$([f]_\beta^\gamma) =$ *column-rank*$([f]_\beta^\gamma)$

(b) $\mathrm{nullity}(f) = \dim(V) - \mathrm{rank}(f)$

Theorem 2.1.5. *Let $f : V \to V$ be a linear map with $\dim V < \infty$. Then the following are equivalent:*

(a) f is 1-1

(b) $N(f) = 0$

(c) f is onto

(d) $\mathrm{rank}(f) = \dim V$

(e) f is an isomorphism.

Why is this true?

Here are some questions you should be able to answer.

- If $f : V \to W$ is linear and invertible, must $f^{-1} : W \to V$ be linear?
- If V, W are finite dimensional and $f : V \to W$ is an invertible linear map, and β, γ bases of V, W, must $[f]_\beta^\gamma$ be invertible? Does your answer depend on the choice of basis β, γ? If the matrix is invertible must f be invertible?

- If V is an n-dimensional vector space over F, is $V \cong F^n$? (here "$\cong$" means "isomorphic to").

You should know how to work with change of coordinate matrices.

Definition 2.1.6. If V is a vector space, β, β' two bases for V, then the change of coordinate matrix associated to β, β' is

$$[I_V]_{\beta'}^{\beta}$$

Theorem 2.1.7. *With notations as in the definition, if $v \in V$ is any vector, then*

$$[I_V]_{\beta'}^{\beta}[v]_{\beta'} = [v]_{\beta}$$

Moreover, if $T : V \to V$ is a linear map, then

$$[T]_{\beta'} = ([I_V]_{\beta'}^{\beta})^{-1}[T]_{\beta}[I_V]_{\beta'}^{\beta}$$

You should know how to prove this. You should know how to use this to express a linear transformation as a matrix with respect to different bases. Some more questions you should be able to answer:

(a) If A is a matrix, what is Ae_i?

(b) If $T : V \to V$ and β a finite basis for V , then what is $[T]_{\beta}e_i$? What is $[T]_{\beta}[\beta_i]_{\beta}$?

Definition 2.1.8. Two matrices $A, B \in M_n(F)$ are similar if there exists an invertible Q such that $A = Q^{-1}BQ$.

- Is every invertible matrix Q a change of coordinate matrix?
- If β, β' are unequal bases of a vector space, could $[T]_{\beta} = [T]_{\beta'}$? (The answer is yes. How?)

2.2 Determinants

Definition 2.2.1. The determinant of $A \in M_n(F)$ is

$$\det A := \sum_{j=1}^{n}(-1)^{1+j}A_{1j}\det \widetilde{A_{1j}}$$

This expression is the "cofactor expansion along the 1st row". This is an element of the base field F.

You can equivalently cofactor expand along any row or column

Theorem 2.2.2. *For $A \in M_n(F)$, for any $1 \le i \le n$, we have*

$$\det A = \sum_{j=1}^{n}(-1)^{i+j}A_{ij}\det \widetilde{A_{ij}} = \sum_{j=1}^{n}(-1)^{i+j}A_{ji}\det \widetilde{A_{ji}}$$

You should know that determinants detect invertibility:

Theorem 2.2.3. *If $A \in M_n(F)$, then A is invertible if and only if* $\det A \neq 0$.

You should know that determinants are multiplicative:

Theorem 2.2.4. *If $A, B \in M_n(F)$, then* $\det(AB) = \det A \cdot \det B$.

How did we prove this? This multiplicativity implies a number of properties:

Theorem 2.2.5. *If $A \in M_n(F)$ is invertible and $B \in M_n(F)$, then*

$$\det(A^{-1}) = \det(A)^{-1}$$

and

$$\det(ABA^{-1}) = \det B$$

I.e., similar matrices have the same determinant.

You should know how to prove this from the multiplicativity. Some more questions you should be able to answer:

- How is the determinant affected by elementary row operations?
- What is the determinant of an upper triangular matrix? A lower triangular matrix? A diagonal matrix?
- What is the determinant of a linear operator $T : V \to V$? Does it depend on a basis?

You should know how to evaluate determinants efficiently either using cofactor expansion or elementary row operations.

2.3 Eigenvalues, Eigenvectors, Eigenspaces, Eigenbases

Definition 2.3.1. Let $T : V \to V$ be linear. A scalar $\lambda \in F$ is an eigenvalue of T if there is a nonzero vector v such that $T(v) = \lambda v$. Such a vector v is an eigenvector for the eigenvalue λ (or a λ-eigenvector). For an eigenvalue λ of T, the λ-eigenspace E_λ of T is the subspace of V spanned by the λ-eigenvectors. It can also be described as:

$$E_\lambda = \{v \in V \mid T(v) = \lambda v\} = N(T - \lambda I)$$

You should know why the equalities above are true. An eigenbasis for V is a basis of V consisting of eigenvectors of T.

Definition 2.3.2. Let V be a finite dimensional vector space. A linear map $T : V \to V$ is diagonalizable if there is a basis β for V such that $[T]_\beta$ is diagonal.

Theorem 2.3.3. *Let T be a linear operator on a finite dimensional vector space V. Then T is diagonalizable if and only if V has an eigenbasis.*

How did we prove this?

- What are the eigenvalues of a diagonal matrix?
- Do diagonal matrices have an eigenbasis? If so, what is it?
- Are diagonal matrices invertible?
- Is the characteristic polynomial of an upper triangular matrix split?
- Is the characteristic polynomial of a diagonalizable matrix split?

Definition 2.3.4. The characteristic polynomial of a matrix $A \in M_n(F)$ is the polynomial

$$\chi_A(t) := \det(A - tI_n)$$

Why is $\det(A - tI_n)$ a polynomial? What is its degree? Its leading coefficient? Its constant term? If Q is invertible, is $\chi_{Q^{-1}AQ} = \chi_A$? Why?

Theorem 2.3.5. *The eigenvalues of A are the roots of $\chi_A(t)$.*

- Why are the eigenvalues of A exactly the roots of $\chi_A(t)$?
- Can 0 be an eigenvalue?
- How many eigenvalues can a matrix in $M_n(F)$ have? Can a matrix have no eigenvalues?
- For an eigenvalue λ, how do you compute a basis for the λ-eigenspace?
- For distinct eigenvalues μ, λ, what is $\dim E_\mu \cap E_\lambda$?
- For distinct eigenvalues μ, λ, if $v_1, \ldots, v_r$ are linearly independent μ-eigenvectors and $w_1, \ldots, w_s$ are linearly independent λ-eigenvectors, could $\{v_1, \ldots, v_r, w_1, \ldots, w_s\}$ be linearly dependent?
- For a root λ of $\chi_A(t)$ with multiplicity m_λ, why is $1 \leq \dim E_\lambda \leq m_\lambda$?
- What is the analogous statements/questions/answers to the above for linear operators instead of matrices? Can you give examples illustrating your answers to the above questions?

Theorem 2.3.6. *A linear operator T on a finite dimensional vector space V is diagonalizable if and only if $\chi_T(t)$ splits and for every eigenvalue λ, $\dim E_\lambda$ equals the multiplicity of λ.*

You should be able to describe the possibilities for the number and multiplicities of the eigenvalues and the dimensions of the eigenspaces, giving examples as appropriate, as done on homework 8. Can you do this in the case $\dim V = 4$?

You should be able to check a matrix A for diagonalizability, and if diagonalizable, find an eigenbasis and a matrix Q such that $Q^{-1}AQ$ is diagonal. Is the matrix Q unique? Why or why not? You should also be able to do this for matrices over $\mathbb{C}$.

2.4 Invariant subspaces and Cayley-Hamilton

Definition 2.4.1. Let $T : V \to V$ be linear. A subspace $W \subset V$ is T-invariant if $T(W) \subset W$.

- What are some examples of T-invariant subspaces?
- If $W \subset V$ is T-invariant, must $T(W) = W$? Is W also T^2-invariant? Is W also $(T^2 + T)$-invariant? Is W also $(T^2 - \lambda I)$-invariant? (for any scalar $\lambda \in F$)
- Is the sum of T-invariant subspaces T-invariant? Is a subspace of a T-invariant subspace T-invariant?
- If $v \in V$ is a vector, what is a basis for the T-invariant subspace generated by v? (We denoted this by $\langle v \rangle_T$ in class. What is the matrix of T with respect to this basis?
- If $W \subset V$ is T-invariant, what is the characteristic polynomial of T_W and what is its relationship to χ_T?

Theorem 2.4.2 (Cayley-Hamilton theorem)**.** *If T is a linear operator on a finite dimensional vector space V, then $\chi_T(T) = 0$.*

- What does this theorem mean? How do you prove it?
- Consider the fake proof: "$\chi_T(T) = \det(T - TI) = \det(T - T) = \det(0) = 0$". Why does this proof not make sense?
- Let $A \in M_n(F)$ be invertible. How can you use the Cayley-Hamilton theorem to find a formula for A^{-1} in terms of powers of A?

2.5 Generalized eigenvectors, generalized eigenspaces

Definition 2.5.1. Let $T : V \to V$ be a linear operator. If λ is an eigenvalue of T, then $v \in V$ is a generalized λ-eigenvector of T if $v \neq 0$ and $(T - \lambda I)^p v = 0$ for some positive integer p. The generalized λ-eigenspace is

$$K_\lambda := \{v \in V \mid (T - \lambda I)^p v = 0 \text{ for some positive integer } p.\}$$

Are generalized eigenspaces T-invariant? Why? What is its relation to eigenvectors/eigenspaces?

3 Final Exam Review

Compared to the previous exams, the additional material on the final consists mostly of sections 7.1, 7.2, and whatever we cover on Friday December 9. In this review, we pose some questions, whose answers can be found at the end of these notes. You can use these questions as practice problems for the final. I will continue to add to and edit these notes as we approach the final exam date. As usual, anyone who finds a mathematical error in these notes will be awarded a bonus exam point.

Note that these notes *are not a replacement* for the book or lecture notes. They are designed to complement the book and lecture notes. The book remains the ultimate reference. Here we just highlight some of the main points, without proofs. As usual, you should at least have a general idea of how to prove most of the Theorem statements shown below. For about 75% of them you should be able to prove on the exam. For all of them you should be able to recover the proof given sufficient time.

For this section, T denotes a linear operator on a finite dimensional vector space V, λ is an eigenvalue of T with multiplicity m_λ, and K_λ is the λ-generalized eigenspace, defined to be

$$K_\lambda = \bigcup_{p\geq 1} N((T-\lambda I)^p)$$

Note that $N(T-\lambda I) \subset N((T-\lambda I)^2) \subset N((T-\lambda I)^3) \subset \cdots$.

3.1 Decomposing V as a direct sum of generalized eigenspaces

The first part of the theory of the Jordan canonical form shows that we can decompose V into a direct sum of generalized eigenspaces. Here are some statements.

Theorem 3.1.1. *The following are true:*

(a) *K_λ is a T-invariant subspace containing E_λ. (Theorem 7.1 in the book)*

(b) *$K_\lambda = N((T-\lambda I)^{m_\lambda})$. (Theorem 7.2(b) in the book)*

(c) *For any two distinct scalars μ, λ, $K_\lambda \cap K_\mu = 0$, and $(T-\mu I)_{K_\lambda}$ is an isomorphism $K_\lambda \to K_\lambda$. (Theorem 7.1 in the book)*

(d) *If χ_T splits and $\lambda_1, \ldots, \lambda_k$ are the eigenvalues of T, then for any $v \in V$, there exist unique $v_i \in K_{\lambda_i}$ such that $v = v_1 + \cdots + v_k$. (Theorem 7.3 in the book) Moreover, with the same assumptions, if β_i is a basis for K_{λ_i}, then $\beta_1 \cup \beta_2 \cup \cdots \cup \beta_k$ is a basis for V, and $\dim K_{\lambda_i}$ is equal to the multiplicity of λ_i.*

(e) *If χ_T splits, then V is the direct sum of its generalized eigenspaces. (Theorem 7.8 in the book)*

Questions

Q1a Show that T is diagonalizable if and only if χ_T is split and $E_\lambda = K_\lambda$ for each eigenvalue λ.

Q1b Suppose $W \subset V$ is a T-invariant subspace. Show that $W \cap E_\lambda$ and $W \cap K_\lambda$ are precisely the λ-eigenspace and λ-generalized eigenspace of T_W respectively.

Q1c Suppose T is diagonalizable, then must T_W be diagonalizable? Why?

Q1d If χ_T is split, must every eigenvalue of T be an eigenvalue of T_W?

Q2 If χ_T splits and $\lambda_1, \ldots, \lambda_k$ are the eigenvalues of T, can you describe a map $P : V \to V$ such that $P(K_{\lambda_1}) = K_{\lambda_1}$ but $P(K_{\lambda_i}) = 0$ for $i > 1$? To make things more concrete, you can consider the matrix

$$T = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 1 & 2 & 3 & 4 \\ 0 & 0 & 2 & 3 & 4 \\ 0 & 0 & 0 & 2 & 3 \\ 0 & 0 & 0 & 0 & 3 \end{bmatrix}$$

with characteristic polynomial $\chi_T(t) = (1-t)^2(2-t)^2(3-t)$. Can you describe a map $P : \mathbb{R}^5 \to \mathbb{R}^5$ such that $P(K_1) = K_1$ but $P(K_2) = P(K_3) = 0$? Hint: Consider a certain polynomial in T. Think about parts (b) and (c) of the Theorem.

Q3a A linear operator $U : V \to V$ is *quasi-unipotent* if $U^p = I$ for some $p \geq 1$. Some examples are: $U = I, U = \left[\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right], U = \left[\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right]$. Let λ be an eigenvalue of U. Show that $\lambda^n = 1$ for some positive integer n (i.e., λ is a "root of unity").

Q3b Without multiplying it out, show that the matrix

$$U = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

is quasi-unipotent, and find the minimum $p \geq 1$ such that $U^p = I$. What are its eigenvalues over $\mathbb{C}$? Find bases for the eigenspaces E_1 and E_{-1}. Hint: Consider how U acts on the standard basis.

3.2 Decomposing K_λ into a direct sum of spans of cycles

With notation as in Theorem 3.1.1(d), part (d) implies that $[T]_\beta$ is block diagonal, with blocks $[T_{K_{\lambda_i}}]_{\beta_i}$ (for $i = 1, 2, \ldots, k$). The next task is to describe a good basis for each K_{λ_i} so that the blocks $[T_{K_{\lambda_i}}]_{\beta_i}$ are relatively simple. For this, we will need the notion of a cycle of generalized eigenvectors.

Definition 3.2.1. Let x be a generalized λ-eigenvector (for T), and let $p \geq 1$ be the minimum positive integer such that $(T - \lambda I)^p x = 0$. Then

$$C_x := \{(T - \lambda I)^{p-1}x, (T - \lambda I)^{p-2}x, \ldots, (T - \lambda I)x, x\}$$

is called the *cycle of generalized eigenvectors generated by x*. Here, $(T - \lambda I)^{p-1}x$ is called the *initial vector* of the cycle, and x is called the *end vector* (or generating vector) of the cycle. The length of the cycle C_x is the number of elements in C_x, namely, p. Note that for a cycle of length 1, the initial vector is the same as the end vector.

Theorem 3.2.2. *For a cycle C, the initial vector is the only eigenvector in the cycle, and if x is an eigenvector then $C_x = \{x\}$ has length 1.*

A key question to consider is: Is C_x linearly independent? Does V have a basis consisting of cycles? It turns out the answer to both questions is yes:

Theorem 3.2.3. *As usual let $T : V \to V$ be a linear operator on a finite dimensional vector space V, and let λ be an eigenvalue of T.*

(a) (Theorem 7.6 in the book) If $C_1, \ldots, C_n$ are cycles of generalized λ-eigenvectors such that the initial vectors are distinct and linearly independent, then the C_i's are disjoint and $\cup_{i=1}^n C_i$ is linearly independent.

(b) (Theorem 7.7 in the book) K_λ admits a basis consisting of cycles of generalized eigenvectors.

As a corollary of (a), we find that every cycle is itself linearly independent. Moreover, if C_x is a cycle of length p generated by x, and if we write $C_{x,i} := (T - \lambda I)^i x$, then we have (for $0 \leq i \leq p - 1$)

$$(T - \lambda I)(T - \lambda I)^i x = (T - \lambda I)C_{x,i} = (T - \lambda I)^{i+1}x = C_{x,i+1} \tag{1}$$

where we set $C_{x,p} = 0$. On the other hand, we also have:

$$(T - \lambda I)(T - \lambda I)^i x = T(T - \lambda I)^i x - \lambda I(T - \lambda I)^i x = TC_{x,i} - \lambda C_{x,i} \tag{2}$$

Putting (1) and (2) together, we find that

$$TC_{x,i} - \lambda C_{x,i} = C_{x,i+1}, \qquad \text{or equivalently} \qquad TC_{x,i} = \lambda C_{x,i} + C_{x,i+1}$$

It follows that if we order the vectors in C_x from initial vector to end vector, then

$$[T_{\langle C_x \rangle}]_{C_x} = \begin{bmatrix} \lambda & 1 & & & & \\ & \lambda & 1 & & & \\ & & \lambda & 1 & & \\ & & & \ddots & \ddots & \\ & & & & \lambda & 1 \\ & & & & & \lambda \end{bmatrix} \tag{3}$$

where the missing entries are all 0. By part (b), if $\beta_\lambda = C_{x_1} \cup C_{x_2} \cup \cdots \cup C_{x_r}$ is a basis for K_λ consisting of cycles, then $[T_{K_\lambda}]_\beta$ is block diagonal with each block of the form (3). Such a block is called a *Jordan block* of T.

Theorem 3.2.4 (Existence of the Jordan Canonical Form). *If χ_T is split, and $\lambda_1, \ldots, \lambda_k$ are the distinct eigenvalues of T, and β_{λ_i} is a basis of K_{λ_i} consisting of a union of cycles of generalized λ_i-eigenvectors, then $\beta := \beta_{\lambda_1} \cup \cdots \cup \beta_{\lambda_k}$ is a basis of V, and $[T]_\beta$ is block diagonal, with each block of the form* (3).

Can you prove this from the facts discussed above?

In the theorem, if each β_{λ_i} is arranged so that the cycles appear in order of decreasing length, and within each cycle the initial vector appears first and end vector appears last, then β is called a *Jordan canonical basis*, and the matrix $[T]_\beta$ is called a *Jordan canonical form* of T.

Questions.

Q4a Suppose $V = F^n$ and T is a linear operator on V with split characteristic polynomial. Let J be a Jordan canonical form of T, viewed as a linear operator on $V = F^n$. Is T similar to J?

Q4b If $A \in M_n(F)$, viewed as a linear operator on $V = F^n$, then we say that A is upper-triangularizable if it is similar to an upper-triangular matrix. Is it true that A is upper-triangularizable if and only if χ_A splits? Prove your answer. (Hint: Use Q4a)

3.3 Dot diagrams and computing Jordan blocks

Here we focus on T_{K_λ}. The existence of a basis β of K_λ consisting of cycles (i.e., a Jordan canonical basis for T_{K_λ} allows us to describe the possible forms of the matrix $[T_{K_\lambda}]_\beta$. If $\beta = C_{x_1} \cup C_{x_2} \cup \cdots \cup C_{x_r}$ for various generalized eigenvectors $x_1, \ldots, x_r$, then we will often arrange them in a "dot diagram" (see the book for details). As an example, let $U := (T - \lambda I)_{K_\lambda}$, and suppose

$$\begin{aligned}
C_{v_1} &= \{U^4 v_1, U^3 v_1, U^2 v_1, U v_1, v_1\} \\
C_{v_2} &= \{U^4 v_1, U^3 v_2, U^2 v_2, U v_2, v_2\} \\
C_{v_3} &= \{U^2 v_3, U v_3, v_3\} \\
C_{v_4} &= \{U v_4, v_4\} \\
C_{v_5} &= \{U v_5, v_5\} \\
C_{v_6} &= \{v_6\}
\end{aligned}$$

are cycles such that $\beta := \cup_{i=1}^{6} C_{v_i}$ is a basis for K_λ. Then we will often arrange the vectors in an diagram as follows:

$$\begin{array}{cccccc}
U^4 v_1 & U^4 v_2 & U^2 v_3 & U v_4 & U v_5 & v_6 \\
U^3 v_1 & U^3 v_2 & U v_3 & v_4 & v_5 & \\
U^2 v_1 & U^2 v_2 & v_3 & & & \\
U v_1 & U v_2 & & & & \\
v_1 & v_1 & & & &
\end{array} \tag{4}$$

The corresponding "dot diagram" would be

$$\begin{array}{cccccc}
\bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\
\bullet & \bullet & \bullet & \bullet & \bullet & \\
\bullet & \bullet & \bullet & & & \\
\bullet & \bullet & & & & \\
\bullet & \bullet & & & &
\end{array} \tag{5}$$

where each dot "$\bullet$" represents a vector in (4). Thus, the cycles are arranged as columns, with initial vectors on top and end vectors at the bottom. The cycles are arranged from longest (at the left) to shortest (at the right). Of course, this arrangement is not unique, since you could have swapped the first two columns in (4). However, note that this does not change the dot diagram, so the dot diagram is uniquely determined from T using these "arrangement rules".

Note that $U = (T - \lambda I)_{K_\lambda}$ sends each vector in (4) to the vector above it. For vectors in the first row, U sends them to 0. Thus it follows that nullity(U) is at least the size of the first row (equivalently, the number of cycles in β). Similarly, for $i \geq 1$, U^i sends anything in the first i rows to 0, and sends everything else to the vector i positions above it.

Questions. Easy version: prove these for the special case described above. Slightly harder version: prove them in general.

- Q5a For $i \geq 1$, let $\beta' \subset \beta$ be the subset consisting of vectors which have at least i vectors below it in the diagram. Prove that β' is a basis for $R(U^i)$.
- Q5b For $i \geq 1$, prove that the first i rows is a basis for $N(U^i)$. Hint: Use Q5a.

Summarizing the results of Q5a and Q5b above, we get

Theorem 3.3.1 (compare with Theorem 7.9 in the book)**.** *For any $i \geq 1$, the first i rows of the diagram* (4) *are a basis for $N(U^i)$, and the complement of the first i rows*[1]*, shifted up by i, are a basis for $R(U^i)$.*

- Q6 Could a vector in β simultaneously lie in $N(U^i)$ *and* $R(U^i)$? Is this weird?

Theorem 3.3.2 (See the Corollary on p494 in the book)**.** *Suppose χ_T is split with distinct eigenvalues $\lambda_1, \dots, \lambda_k$. Then for each eigenvalue λ_i the Jordan canonical form of $T_{K_{\lambda_i}}$ is uniquely determined by its dot diagram. Thus, the Jordan canonical form of T is uniquely determined up to an ordering of its eigenvalues.*

This statement holds because we have asserted that for a basis of cycles to be a Jordan canonical basis, the cycles must appear in order of descending length. The choice to impose such a restriction is mostly for aesthetic reasons. For a different ordering of the cycles, the associated ordered basis will still give rise to a block diagonal matrix with blocks of the form (3).

3.4 The minimal polynomial

Definition 3.4.1. Let $T : V \to V$ be linear as usual with V finite dimensional. A polynomial $p(t)$ is a minimal polynomial for T if it is monic, $p(T) = 0$ (the zero operator), and $p(t)$ is of lowest positive degree amongst monic polynomials satisfying $p(T) = 0$.

Theorem 3.4.2 (Theorem 7.12 in the book)**.** *Let $p(t)$ be a minimal polynomial for T. If $g(t)$ is any polynomial such that $g(T) = 0$, then $p(t)$ divides $g(t)$. Moreover, the minimal polynomial is unique.*

A key point is that the minimal polynomial divides the characteristic polynomial. This makes computing the minimal polynomial relatively simple in practice.

Questions

- Q7 Let $T : V \to V$ be a linear operator with split characteristic polynomial and exactly two eigenvalues $\lambda_1 = 1$ and $\lambda_2 = 2$. Suppose the dot diagrams for λ_1, λ_2 are, respectively,

$$\lambda_1 = 1 : \begin{array}{ccc} \bullet & \bullet & \bullet \\ \bullet & & \end{array} \qquad , \qquad \lambda_2 = 2 : \begin{array}{cc} \bullet & \bullet \\ \bullet & \bullet \\ \bullet & \end{array}$$

  What are the dimensions of E_1, E_2, K_1, K_2? What is $\chi_T(t)$? What is minpoly$_T(t)$? What is the Jordan canonical form of T?

Additional resources for studying. In addition to these practice problems, I would also encourage you to work through some of the examples and exercises from sections 7.1, 7.2, and 7.3 in the textbook. Solutions to many of the exercise problems involving finding Jordan canonical forms can be found via the free online tool wolframalpha.

[1] This means $\beta - \{\text{first } i \text{ rows}\}$.

4 Answers to questions posed in Final Exam Review

Q1a **Solution.** We know T is diagonalizable if and only if χ_T is split and $\dim E_\lambda = m_\lambda = \dim K_\lambda$. But since $E_\lambda \subset K_\lambda$, they have the same dimension if and only if they are equal. This is a corollary in §7.1 in the book.

Q1b **Solution.** For $w \in W$, w lies in the λ-generalized eigenspace of T_W if and only if $(T_W - \lambda I)^p w = 0$ for some $p \geq 1$. Since $T_W w = Tw$, it follows that $(T - \lambda I)^p w = 0$, so $w \in K_\lambda$, so $w \in W \cap K_\lambda$. Conversely, suppose $w \in W \cap K_\lambda$. Then since $w \in K_\lambda$, $(T - \lambda I)^p w = 0$ for some $p \geq 1$, but again since $Tw = T_W w$, this means that $(T_W - \lambda I)^p w = 0$, so w lies in the λ-generalized eigenspace of T_W. The same argument shows that $W \cap E_\lambda$ is the λ-eigenspace of T_W.

Q1c **Solution.** Yes. T is diagonalizable if and only if χ_T is split and $K_\lambda = E_\lambda$. Since $W \cap E_\lambda, W \cap K_\lambda$ are the eigenspaces and generalized eigenspaces of T_W, we also have that T_W is diagonalizable if and only if χ_{T_W} is split and $W \cap K_\lambda = W \cap E_\lambda$ for each eigenvalue λ of T_W. Since T is diagonalizable, χ_T is split, and since χ_{T_W} divides χ_T, χ_{T_W} is also split. Since $K_\lambda = E_\lambda$, it follows immediately that $W \cap K_\lambda = W \cap E_\lambda$. Thus T_W is diagonalizable.

Q1d **Solution.** No. For example, take $T = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 2 \end{smallmatrix}\right]$, and $W = \mathrm{Span}\{e_1\}$. Then $\chi_T = (1-t)(2-t)$, but 2 is an eigenvalue of T which is not an eigenvalue of T_W.

Q2 **Solution.** You can take $P = (T - \lambda_2 I)^{m_{\lambda_2}} \cdots (T - \lambda_k I)^{m_{\lambda_k}}$. In the specific example, you can take $P = (T - 2I)^2(T - 3I)$. Think about why this works.

Q3a **Solution.** Suppose λ is an eigenvalue, then there is a nonzero vector $v \in V$ such that $Uv = \lambda v$. Since U is quasi-unipotent, there is some $p \geq 1$ such that $U^p = I$, so $U^p v = v$. But we also have $U^p v = (U^{p-1})Uv = U^{p-1}\lambda v = U^{p-2}\lambda^2 v = \cdots = \lambda^p v$. This shows that $v = \lambda^p v$, so $0 = v - \lambda^p v = (1 - \lambda^p)v$. Since $v \neq 0$, this shows that $1 - \lambda^p = 0$, i.e., $\lambda^p = 1$.

Q3b **Solution.** Note that $Ue_1 = e_2$, $Ue_2 = e_3$, $Ue_3 = e_4$, and $Ue_4 = e_1$. This shows that $U^4 = I$, and 4 is the smallest positive integer for which this holds. The characteristic polynomial can be computed easily using cofactor expansion along the top row. Doing this yields $\chi_U(t) = t^4 - 1 = (t-1)(t+1)(t^2+1)$. Thus its eigenvalues are $1, -1, i, -i$, and hence is diagonalizable over $\mathbb{C}$ with 1-dimensional eigenspaces. One easily checks that $(1,1,1,1) \in E_1$, so it is a basis, and similarly $(1,-1,1,-1) \in E_{-1}$, so it is a basis for E_{-1}.

Q4a **Solution.** Yes. Let $\beta = \{\beta_1, \beta_2, \ldots, \beta_n\}$ be a Jordan canonical basis for T. If Q is the linear operator on $V = F^n$ sending e_i to β_i, then $T = Q^1 JQ$, so T is similar to J. To check this, it suffices to check this on a basis. Thus, you should check that indeed

$$Te_i = Q^{-1}JQe_i \qquad \text{for each } i \in \{1, \ldots, n\}$$

Q4b **Solution.** Yes. Suppose A is upper-triangularizable. Then there is an upper triangular matrix U which is similar to A, so $\chi_A = \chi_U$, but $\chi_U = (U_{11} - t)(U_{22} - t) \cdots (U_{nn} - t)$, where U_{ij} denotes the (i,j)-th entry of U. This shows that $\chi_A = \chi_U$ is split. Conversely, if χ_A is split, then A is similar to its Jordan canonical form J, which is upper triangular.

Q5a **Solution.** Since β spans K_λ (the domain of U^i),

$$R(U^i) = \mathrm{Span}\{U^i(v) \mid v \in \beta\}$$

Since U^i moves each vector up by i positions and sends the first i rows to 0, $\{U^i(v) \mid v \in \beta\}$ consists of exactly the vectors in β which lie above at least i vectors in the diagram. I.e., $\{U^i(v) \mid v \in \beta\} = \beta'$. Thus β' spans $R(U^i)$. On the other hand, β' is a subset of β, and since β is linearly independent, so is β', so β' is a basis for $R(U^i)$.

Q5b **Solution.** Clearly the first i rows are contained in $N(U^i)$. Again since they are a subset of β and β is linearly independent, they are also linearly independent. To show that they are a basis for $N(U^i)$, it remains to show that they span $N(U^i)$. Since $N(U^i)$ is finite dimensional (since it is a subspace of the finite dimensional space V), it suffices to check that the first i rows has exactly $\dim N(U^i)$ elements. From the rank-nullity theorem, we have

$$|\beta| = \dim K_\lambda = \operatorname{rank}(U^i) + \operatorname{nullity}(U^i)$$

By Q5a, we know that $\operatorname{rank}(U^i)$ is exactly the number of vectors which have at least i vectors below them. By shifting these vectors down i positions, this is also exactly the number of vectors which are not in the first i rows. Thus $\operatorname{rank}(U^i) = |\beta| -$ (the number of vectors in the first i rows). It follows that

$$\operatorname{nullity}(U^i) = \dim N(U^i) = \text{the number of vectors in the first } i \text{ rows}$$

This completes the proof.

Q6 **Solution.** Yes! In our example (4), we in fact have $R(U^3) \subset N(U^3)$. What this says is that $(U^3)^2 = U^6 = 0$ (applying U^3 once lands in $N(U^3)$, so applying it again gives you 0). This is not weird at all. For example, the matrix $A := \left[\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right]$ satisfies $R(A) \subset N(A)$, since $A^2 = 0$. A key point to remember is that even though $R(U^i)$ and $N(U^i)$ are in this case both subspaces of the same vector space K_λ, in general the former is a subspace of the codomain (of U^i), and the latter is a subspace of the domain (of U^i).

Q7 **Solution.** $\dim E_1 = 3, \dim E_2 = 2, \dim K_1 = 4, \dim K_2 = 5$, $\chi_T(t) = (1-t)^4(2-t)^5$, and $\operatorname{minpoly}_T(t) = (1-t)^2(2-t)^3$. Note that since $\operatorname{minpoly}_T$ divides χ_T, there aren't many options for $\operatorname{minpoly}_T(t)$. You should convince yourself that no factor $f(t)$ of $(1-t)^2(2-t)^3$ can possibly satisfy $f(T) = 0$.